

## FluenCi and the Dynamic Speech Corpus (DSC)

**Dermot F. Campbell, Marty Meinardi, Charlie Pritchard, Yi Wang**

Dublin Institute of Technology (Ireland)

[dermotfcampbell@gmail.com](mailto:dermotfcampbell@gmail.com)

### Abstract

The EU funded FluenCi project aims to introduce learners to the importance of intonation in native-to-native communication via high-frequency collocations. Students are presented with common lexical and syntactical structures in neutral form plus several intonational variations, illustrating different speaker intentions and realised by several speech features such as stress, vowel length, pitch change and speed of delivery.

The terminology used to describe the scripted audio assets in the FluenCi database is the same as that used in DIT's Dynamic Speech Corpus, which is a corpus of native-to-native dialogues designed primarily for language learners. This corpus sensitises users to dialogue as response, rather than simply speech production, stressing the interactive strategies of speakers and making the resultant phonetic realisations accessible to learners and researchers alike.

The DSC framework is based on a unique approach to natural, informal speech and embodies this hierarchical analysis of speech in an XML database with outputs based on an innovative dialogue player where the speech and transcripts are automatically synchronised, regardless of playback speed.

The DSC can also accommodate various dialogue categories such as scripted, semi-scripted, un-scripted and re-scripted, providing a full range of naturalness ranging from materials which can be used to teach vocabulary and structure to natural dialogues which capture – and make available for study – the 'messiness' of natural speech (Cauldwell 2002) in a principled fashion, including the use of a slow-down algorithm.

Whereas the natural dialogues of the DSC offer all learners an accessible model of informal native dialogue, the scripted dialogues of the FluenCi sub-programme will provide a useful scaffolding bridge between the teaching dialogues of currently available language materials and the unscripted audio assets of the DSC. Both projects are linguistically driven and prototypes are currently being developed using HTML5, which will facilitate implementation across desktop, tablet and smartphone.